A MODIFICATION OF LOCAL ESTIMATION OF LIDAR SIGNAL BY MONTE CARLO METHOD WITH SIMULTANEOUS SIMULATION OF FORWARD AND BACKWARD PHOTON TRAJECTORIES

V.A. Korshunov

Scientific Production Union "Taifun", Obninsk

Received May 27, 1994

A modification of local estimation of lidar signals is proposed which increases the efficiency of Monte Carlo method for the small–angle reception at small distances from a scattering volume. A specific feature of the modification is simultaneous simulation of the forward and backward trajectories, the final estimate being weighted–average over all the trajectories constructed.

Many problems of laser sensing of turbid atmosphere are solved by employing Monte Carlo method.[1] Due to specific boundary conditions imposed (point and pencil–beam sources and receivers), such algorithms are based on local estimations.

In the case when the lidar and the medium sensed are spatially separated and the parameters $\tau_0 = R_0\sigma$ and $\tau_d = \tau_0\varphi$ (with R_0 is the distance between the lidar and the scattering volume, σ is the characteristic value of the layer extinction coefficient, 2φ is the receiving angle) satisfy the conditions $\tau_0 \gtrsim 1$ and $\tau_d \gtrsim 0.1$, the ordinary (single) local estimation is adequate. However, for smaller τ_0 and τ_d the variance of the local estimate sharply increases, thus making it inefficient. The failure is most evident for $\tau_0 = 0$ and vanishing φ, when the double local estimation is used. This is also deficient, in that it slowly converges and involves two (rather than one, as in the local estimation) values of a scattering phase function, thus adding to the spread of the obtained estimates, due to the strong anisotropy of actual phase functions for turbid atmosphere. The existing difficulties have stimulated the development of several modifications to the local estimations[1] some of which we have tried though they gave no satisfactory (at least, as satisfactory as for large τ_0 and τ_d) solution to the problem.

This paper is concerned with the construction of an efficient algorithm for estimation of lidar signals in the case of small τ_0 and τ_d by simulating simultaneously forward and backward photon trajectories. In some aspects, this work develops the approach proposed in Ref. 2.

Consider standard scheme of laser sensing of turbid atmosphere which is characterized by the scattering coefficient $\sigma(\mathbf{r})$, extinction coefficient $\sigma_e(\mathbf{r})$, and the scattering phase function $f(\theta)$ of a unit volume. Let the source of radiation be at a point $\mathbf{a}_s$, and the receiver at $\mathbf{a}_r$. The source directional pattern is described by the function $g_s(\theta)$, while the transmission function of the receiver by $g_r(\theta)$, both satisfying the normalization conditions

$$\int g_s(\theta)\, d\Omega = 1, \quad \int g_r(\theta)\, d\Omega = \Omega_r .$$

Suppose that at the initial moment the source emits a pulse of a unit power of the δ–function time shape. Then the received power P_i averaged over a time interval (t_i, t_{i+1}) can be represented as a series over scattering orders, namely,

$$P_i = \sum_m P_{im} ; \tag{1}$$

$$P_{im} = \Delta\, t_i^{-1}\, \Omega_r\, S_r \int \prod_{k=0}^{m} Q_k(\mathbf{r}_k, \mathbf{r}_{k+1})\, d\mathbf{r}_1 \dots d\mathbf{r}_m, \tag{2}$$

where

$$\mathbf{r}_0 = \mathbf{a}_s,\ \ \mathbf{r}_{m+1} = \mathbf{a}_r,\ Q_k(\mathbf{r}_k, \mathbf{r}_{k+1}) = f_k(\theta_k)\, \sigma(\mathbf{r}_{k+1})\, e^{-\tau_{k,\,k+1}}\, r_{k,\,k+1}^{-2}$$

at $k = 0, 1, \dots, m-1$;

$$Q_m(\mathbf{r}_m, \mathbf{r}_{m+1}) = f_m(\theta_m)\, f_{m+1}(\theta_{m+1})\, e^{-\tau_{m,\,m+1}}\, r_{m,\,m+1}^{-2};$$

$$f_0(\theta_0) = g_s(\theta_0);\ f_{m+1}(\theta_{m+1}) = g_r(\theta_{m+1})\, W_r^{-1},$$

$$f_k(\theta_k) = f(\theta_k) \text{ at } k = 0, 1, \dots, m.$$

In Eq. (2) the sequence of points $(\mathbf{r}_1, \dots, \mathbf{r}_m)$ represents the photon trajectory; θ_k is the scattering angle at $\mathbf{r}_k$; $\mathbf{r}_{k,\,k+1}$ is the distance between the points $\mathbf{r}_k$ and $\mathbf{r}_{k+1}$; $\tau_{k,\,k+1}$ is the optical depth of the path between $\mathbf{r}_k$ and $\mathbf{r}_{k+1}$; S_r is the receiver aperture area; $\Delta t_i = t_{i+1} - t_i$; $\Delta_i = \Delta_i(t_i, t_{i+1}, \mathbf{r}_1, \dots, \mathbf{r}_m) = \theta(t - t_i) - \theta(t - t_{i+1})$, where $\theta(t)$ is the unit step function; $t = \frac{1}{c}\sum_{k=0}^{m} r_{k,\,k+1}$; c is the speed of light. The functions $\theta_k(\mathbf{r}_k, \mathbf{r}_{k+1})$ determine the probability density of transition from $\mathbf{r}_k$ to $\mathbf{r}_{k+1}$.

Let us rearrange Eq. (2) to the form

$$P_{im} = \int R_{0,\,k}(\mathbf{r}_1, \dots, \mathbf{r}_k)\, z_k(\mathbf{r}_{k-1}, \dots, \mathbf{r}_{k+2}) \times$$

$$\times R_{m+1,\,k+1}(\mathbf{r}_{k+1}, \dots, \mathbf{r}_{m+1})\, D_i'\, d\mathbf{r}_1 \dots \qquad d\mathbf{r}_m, \tag{3}$$

where
$k = 0, 1, \dots, m$; $D_i' = \Delta\, t_i^{-1}\, \Omega_r\, S_r\, \Delta_i(t_i, t_{i+1}, \mathbf{r}_1, \dots, \mathbf{r}_m)$;

$$R_{0,\,k} = \prod_{j=0}^{k-1} Q_j(\mathbf{r}_j, \mathbf{r}_{j+1}) \text{ at } k = 1, \dots, m\ ,\ R_{0,0} = 1;$$

$$R_{m+1,\, k+1} = \prod_{j=k+1}^{m} \hat{Q}_j(\mathbf{r}_{j+1}, \mathbf{r}_j);$$

$$\hat{Q}_j(\mathbf{r}_{j+1}, \mathbf{r}_j) = f_{j+1}(Q_{j+1})\, \sigma(r_j)\, \mathrm{e}^{-\tau_{j,\, j+1}}\, r_{j,\, j+1}^{-2}$$

at $k = 0, 1, \dots, m-1$; $R_{m,\, m+1} = 1$;

$$z_k(\mathbf{r}_{k-1}, \dots, \mathbf{r}_{k+2}) = f_k(Q_k)\, f_{k+1}(Q_{k+1})\, \mathrm{e}^{-\tau_{k,\, k+1}}\, r_{k,\, k+1}^{-2}.$$

The quantity $R_{0,\, k}$ in Eq. (3) is the probability density of photon motion from $\mathbf{r}_0$ to $\mathbf{r}_k$, while $R_{m+1,\, k+1}$ denotes the probability density of photon motion along the backward trajectory from $\mathbf{r}_{m+1}$ to $\mathbf{r}_{k+1}$. As follows from Eq. (3), P_{im} can be evaluated by simulating the corresponding segments of forward and backward trajectories and determining the mathematical expectation of $z_k\, \Delta'_i$, that is,

$$P_{im} = \mathrm{M}[z_k\, \Delta'_i]\,,$$

where M denotes the mathematical expectation. Here, k may take any value between 0 and m. We note that $k = m$ corresponds to the ordinary local estimation, while $k = m - 1$ corresponds to the double local estimation.

Now, averaging Eq. (3) over all k between 0 and m and changing the order of summation and integration we have

$$P_{im} = \frac{1}{m+1} \int \Delta'_i \sum_{k=0}^{m} R_{0,\, k}\, z_k\, R_{m+1,\, k+1}\, \mathrm{d}\mathbf{r}_1 \dots \mathrm{d}\mathbf{r}_m. \quad (4)$$

Since the product $R_{0,\, k}\, z_k\, R_{m+1,\, k+1}$ depends only upon the trajectory $(\mathbf{r}_1, \dots, \mathbf{r}_m)$ rather than on the number k, the individual summands can be taken with arbitrary weights $\alpha_k(\mathbf{r}_1, \dots, \mathbf{r}_m)$, which satisfy the condition $\sum_{k=0}^{m} \alpha_k = m + 1$. Substituting α_k into Eq. (4) and restoring the orders of summation and integration yield

$$P_{im} = \frac{1}{m+1} \sum_{k=0}^{m} \int \alpha_k\, R_{0,\, k}\, z_k\, R_{m+1,\, k+1}\, \Delta'_i\, \mathrm{d}\mathbf{r}_1 \dots \mathrm{d}\mathbf{r}_m. \quad (5)$$

Expression (5) determines the estimate of P_{im} of the form

$$P_{im} = \frac{1}{m+1} \sum_{k=0}^{m} \mathrm{M}\,[\alpha_k\, z_k\, \Delta'_i]. \quad (6)$$

Let us choose the weights α_k as follows

$$\begin{aligned} &a_k = 0 && \text{for } z_k > z_{i\mathrm{h}}, \\ &a_k = (m+1)/M_{\mathrm{sum}} && \text{for } z_k \le z_{i\mathrm{h}}, \end{aligned} \quad (7)$$

where M_{sum} is the total number of terms in (6) for which $z_k \le z_{i\mathrm{h}}$, and $z_{i\mathrm{h}}$ is a number depending on the number i of segments in time histogram (how to choose $z_{i\mathrm{h}}$ will be discussed below). Expressions (6) and (7) are a sort of modification of a local estimation, whose variance is finite because $\alpha_k\, z_k$ are bounded from above.

The estimate (6) and (7) differs from ordinary local estimates in that the contribution of one and the same trajectory P_{im} is estimated in $(m + 1)$ different ways of the trajectory simulation, and contribution from each variant to P_{im} is the product of the trajectory $(\mathbf{r}_1, \dots, \mathbf{r}_m)$ occurrence in this variant, proportional to $R_{0,\, k}\, R_{m+1,\, k+1}$ and the value of z_k.

Contributions of all the variants are equal in magnitude, but either have small occurrence and large z_k or vice versa. Clearly, the spread of estimates is mainly due to former variants. Introducing the weights (7), the former variants are excluded, while their contributions to P_{im} are accounted for by increasing the contributions of the latter variants by a factor of $(m + 1)M_s^{-1}$.

The efficiency of estimation proposed depends on the choice of z_{ih}. Obviously, for any i the value of $z_{i\mathrm{h}}$ must correlate with the mean power $P_i = \sum_m P_{im}$ in the ith interval of the time histogram. It is assumed here that $z_{i\mathrm{h}} \sim P_i$, where the proportionality constant is found as follows.

For a given m, the maximum weight α_k occurs for $M_{\mathrm{sum}} = 1$ and equals to $m + 1$ (the case of $M_{\mathrm{sum}} = 0$ will be considered below). In this case the contribution from individual simulated variants to P_i amounts to $\Delta t_i^{-1}\, \Omega_\mathrm{r}\, S_\mathrm{r}\, z_k\, N^{-1}$, where N is the number of trajectories simulated. We require that this particular contribution does not exceed βP_i, where $\beta < 1$ (in practice, β value is chosen from the interval 0.01–0.1). Thus we find

$$z_{i\mathrm{h}} = \beta\, P_i\, N\, \Delta t_i\, (\Omega_\mathrm{r}\, S_\mathrm{r})^{-1}. \quad (8)$$

From Eq. (8), assignment of $z_{i\mathrm{h}}$ requires P_i to be preliminarily specified, e.g., through the preliminary calculation with the proposed scheme and using $z_{i\mathrm{h}}$ values larger than those directly following from Eq. (8).

Of special note is the case of $M_{\mathrm{sum}} = 0$, when for all trajectory segments the inequality $z_k > z_{i\mathrm{h}}$ holds. In this case, we assume that $\alpha_k = 0$ for all k and show below that this introduces negligible biases in the estimates.

To evaluate the contribution of trajectories with $M_{\mathrm{sum}} = 0$, we shall consider the following model situation. Let a homogeneous infinite medium with the extinction coefficient σ contains the source emitting an instantaneous pulse of radiation. We need to determine the probability Q_g that after the dimensionless time $U = \sigma c t$ the photon trajectory containing no segments longer than a given value l_g, with the path traversed after the last collision also included. Obviously, Q_g can be considered as an average estimate of the contribution of trajectories with $M_{\mathrm{sum}} = 0$ over all receiver positions and orientations in space for a time delay U.

A simple estimate of the upper limit on Q_g is at once obtained from $(1 - e^{-l_g})^{N_l}$, where N_l is an integer part of the ratio U/l_g. From this, Q_g vanishes with increasing U or decreasing l_g. However, this value is a crude overestimate. For this reason, Q_g values for most practically useful U values were found by numerical Monte Carlo computations (see Table I).

Shown in the Table I is the probability that photon trajectory after the time U contains no segments longer than l_g.

TABLE I.

l_g	U			
	2	5	10	20
0.5	$3{\cdot}10^{-3}$	$7{\cdot}10^{-8}$	–	–
1.0	0.27	$1.3{\cdot}10^{-2}$	$1.0{\cdot}10^{-4}$	$4{\cdot}10^{-9}$
2.0	–	–	0.17	$2.3{\cdot}10^{-2}$.

From the table we see that Q_g is negligible for $l_g \le 0.5$ at $U \le 10$ and for $l_g \le 1.0$ at $U \ge 10$. Since $z_k \sim e^{-\tau(l_g)}\, l_g^{-2}$, the indicated extremum values of l_g provide an upper limit for z_k and, correspondingly, a lower limit for β, respectively. Thus, the selection of β has to proceed as a compromise between the estimate spread and bias. Such a choice is practically feasible that has been justified by numerical calculations.

Examples of calculations of lidar signal power P are given in Fig. 1 in relative units as functions of dimensionless time U. Triangles indicate data obtained by the method of modified estimation (6), (7). For a comparison, shown by broken lines are calculations using ordinary local estimates. Note that for the computation being illustrated the l_g value varied from 0.05 to 0.4 with $N = 5{\cdot}10^4$ and $\beta = 0.01$.

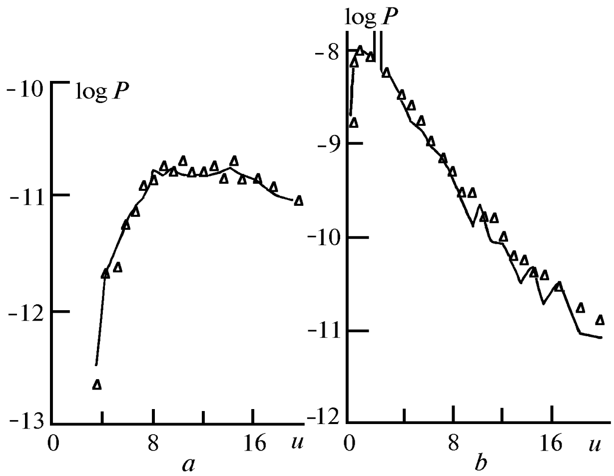

FIG. 1. Signal power vs. dimensionless time; computations using modified estimation (triangles) and local estimations (broken lines): with the lidar and medium spatially separated (a) and with the lidar at the medium boundary (b).

The calculational data in Fig. 1*a* are presented to make a comparison of the developed algorithm with already tested method of local estimation. For variance reduction, at large U the latter was used together with the method of trajectory splitting.[3] The geometry of experiment suggested the lidar and cloud to be spatially separated, with parameters $\tau_0 = 100$ and $\tau_d = 10$. The receiver field of view was formed by a ring diaphragm, so that τ_d referred to the ring outer radius. From Fig. 1*a* it is clear that the two methods agree within the spread of the results.

Let us consider the case in Fig. 1*b*, when both the receiver and the source are located at the boundary of a homogeneous semi–infinite medium, in more detail. The medium extinction coefficient is $\sigma = 20$ km^{-1}, the scattering phase function is that for the Deirmendjian C1 cloud at the wavelength $\lambda = 0.7$ μm, source–receiver separation is 4.2 m. The source has the angular divergence $2\varphi_s = 0.5°$, and its optical axis is normal to the medium boundary. The receiver axis lies in the plane formed by the base segment and the source axis, and is oriented along the direction of 3° away from the source axis. The receiver field of view is $2\varphi_r = 5°$.

The calculations by the method of modified estimation are compared with the double local estimation supplemented by the method of trajectory multiplication. The computation time is the same for both methods. The break in the upper portion of the broken line in Fig. 1*b* corresponds to an order of magnitude spike: a common feature for standard double local estimation techniques.

From Fig. 1*b* it is clear that at small times the two methods agree very well. At $U > 8$ the double local estimation gives systematically low results, with the spread considerably broader than for the modified estimation. To evaluate quantitatively the average spread in results, for $U > 8$ these results were smoothed by least–squares method with two–degree polynomial, the rms deviation δ about the smoothed curve was determined. The δ value for the modified estimation was found to be 0.05, while for the double local estimation it was 0.14, with the bias $\Delta \simeq \delta$. As the time of computing with the method of double local estimation was increased by a factor of 5, δ and Δ were decreased to 0.1. Further calculations to get δ value compared to that in the method of modified estimation were not attempted for the reasons of extraordinary long computer time required. However, the slow convergence of the double local estimation[1] suggests that a twofold reduction of δ would require at least four times longer computer times. Thus, for $U \sim 10$–20, the proposed modification of local estimation is at least one order of magnitude more efficient (in the sense of computation effort required for the same δ) than the double local estimation. We also note that the lidar returns from clouds and fogs can be computed for $U > 20$ using asymptotic relations of transfer theory.[3] Thus, the proposed modification of local estimation allows the Monte Carlo method computation at small τ_0 and τ_d to be done until the coincidence with the results from asymptotic formulas of transfer theory.

REFERENCES

1. G.I. Marchuk, ed., *Monte Carlo Method in Atmospheric Optics* (Nauka, Novosibirsk, 1976), 100 pp.
2. B.M. Golubitskii and M.V. Tantashev, Zh. Vychisl. Mat. Mat. Fiz. **12**, No. 1, 249–252 (1972).
3. V.A. Korshunov, Izv. Vyssh. Uchebn. Zaved. SSSR, Ser. Radiofiz. **30**, No. 10, 1193–1200 (1987).